DOCUMENT RESUME

ED 200 452 SE 034 667

AUTHOR Blosser, Patricia E., Ed.; Mayer, Victor J., Ed.
TITLE Investigations in Science Education, Vol. 7, No. 1. Expanded Abstracts and Critical Analyses of Recent Research.
INSTITUTION ERIC Clearinghouse for Science, Mathematics, and Environmental Education, Columbus, Ohio.; Ohio State Univ., Columbus. Center for Science and Mathematics Education.
PUB DATE 81
NOTE 69p.
AVAILABLE FROM Information Reference Center (ERIC/IRC), The Ohio State Univ., 1200 Chambers Rd., 3rd Floor, Columbus, OH 43212 (subscription $6.00, $1.75 single copy).

EDRS PRICE MF01/PC03 Plus Postage.
DESCRIPTORS *Abstracts; *Mainstreaming; Planetariums; *Science Curriculum; *Science Education; *Science Instruction
IDENTIFIERS Science Education Research

ABSTRACT

Presented are analytical abstracts, prepared by science educators, of research reports in the areas of instruction, mainstreaming, curriculum, classroom learning environment, and the educational use of planetaria. Each abstract includes bibliographical data, research design and procedure, purpose, research rationale, and an abstractor's analysis of the research. Some abstracts are clustered by topics investigated. (CS)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED200452

INVESTIGATIONS IN SCIENCE EDUCATION

U S DEPARTMENT OF HEALTH
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

Editor

Patricia E. Blosser
The Ohio State University

Associate Editor

Victor J. Mayer
The Ohio State University

Advisory Board

Willard J. Jacobson (1982)
Teachers College

Anton E. Lawson (1984)
Arizona State University

Gerald Neufeld (1984)
Brandon University

Livingston S. Schneider (1981)
University of California

Robert L. Steiner (1983)
University of Puget Sound

National Association for Research in Science Teaching

ERIC Clearinghouse for Science, Mathematics
and Environmental Education

Published Quarterly by

The Center for Science and Mathematics Education
College of Education
The Ohio State University
1945 North High Street
Columbus, OH 43210

Subscription Price: $6.00 per year. Single Copy Price: $1.75
Add 50¢ for Canadian mailings and $1.00 for foreign mailings.

SE 034 607

OCT 2

Volume 7, Number 1, 1981

NOTES FROM THE EDITORS. iii

INSTRUCTION 1

Abraham, Michael R. "The Effect of Grouping of Verbal Interaction During Science Inquiries." Journal of Research in Science Teaching, 13(2): 127-135, 1976.
Abstracted by PATRICIA H. SUTER 3

Arlin, Marshall and Ian Westbury. "The Leveling Effect of Teacher Pacing on Science Content Mastery." Journal of Research in Science Teaching, 13(3): 213-219, 1976.
Abstracted by ELIZABETH KEAN. 8

Combs, L. L. "A Teaching Approach to Physical Chemistry." Journal of Research in Science Teaching, 13(5): 467-472, 1976.
Abstracted by J. DUDLEY HERRON. 16

Eccles, P. and G. J. A. Deleeuw, "Information Processing and Pupil Achievement: A Model and A Study." Journal of Research in Science Teaching, 13(5): 383-395, 1976.
Abstracted by WILLIS H. HORAK 22

Lawrenz, Frances. "The Stability of Student Perception of the Classroom Learning Environment." Journal of Research in Science Teaching, 14(1): 77-81, 1977.
Abstracted by CHRIS A. POULER 29

Lloyd, Gerald. "An Experiment with the Teaching of Science to Nonstreamed Classes." The School Science Review, 58: 355-360, 1976.
Abstracted by DOROTHY L. GABEL. 35

Mayfield, J. M. "Factors Affecting Rationality in the Discussions of a Problem by Small Groups of Secondary School Students." Science Education, 60(2): 173-183, 1976.
Martin, J. F. "Analysis of John M. Mayfield's Factors Affecting Rationality in the Discussion of a Problem by Small Groups of Secondary Students." Science Education, 60(2): 185-192, 1976.
Abstracted by DOROTHY L. GABEL. 40

Shymansky, James A. "How Is Student Performance Affected by the One-to-One Teacher-Student Interactions Occurring in an Activity-Centered Science Classroom?" Journal of Research in Science Teaching, 13(3): 253-258, 1976.
Abstracted by WILLIAM R. BROWN. 46

Spears, Jacqueline and Dean Zollman. "The Influence of Structured Versus Unstructured Laboratory on Students' Understanding the Process of Science." Journal of Research in Science Teaching, 14(1): 33-38, 1977.
Abstracted by MARVIN F. WIDEEN. 51

Sunal, Dennis W. "Analysis of Research on the Educational Uses of a Planetarium." Journal of Research in Science Teaching, 13(4): 345-349, 1976.
Abstracted by GERALD H. KROCKOVER 57

RESPONSE TO ANALYSIS. 61

Mills, Terence and Eubanks, I. Dwaine. "Student Assessment of ISCS Classrooms and Grade Placement of ISCS Level I and II," by Dorothy Gabel. Investigations in Science Education, 5(4): 19-22, 1979.
Response by TERENCE MILLS 63

NOTES FROM THE EDITOR

This first issue of Volume 7 contains ten papers related to eleven articles all having some emphasis on instruction. One paper combines two articles, one by Mayfield and the other by Martin, because Martin's article involved a critical review of Mayfield's research. Science content and the educational level of the students involved varies with the article being considered. Abraham looked at the relationship of high school chemistry instruction and divergent thinking. Lawrenz described the use of the Learning Environment Inventory for gaining information about student perceptions of their classroom environment. Eccles and Deleeuw looked at student information processing behaviors. The Mayfield and Martin articles are concerned with small group interaction in a problem solving situation while Shymansky looked at teacher-student interaction in activity-based science lessons. Three articles have a common focus on the pace of instruction: Arlin and Westbury, who looked at the influence of teacher control of the pace of instruction on learning rates; Combs, who studied teacher pacing vs. self pacing in physical chemistry; and Lloyd, who examined the use of worksheets in self paced and group paced settings. Spears and Zollman compared the use of structured vs. unstructured laboratories in introductory college physics. Sunal reported on the use of the school planetarium in education.

This issue also contains a response to an article critiqued in a previous issue of ISE. The critique appeared in Volume 5, issue 4.

Patricia E. Blosser
Editor

Victor J. Mayer
Associate Editor

INSTRUCTION

Abraham, Michael R. "The Effect of Grouping on Verbal Interaction During Science Inquiries." Journal of Research in Science Teaching, 13(2): 127-135, March 1976.
Descriptors--*Chemistry; *Divergent Thinking; Educational Research; *Grouping (Instructional Purposes); Instruction; Science Education; Secondary Education; *Secondary School Science; *Verbal Communication

Expanded abstract and analysis prepared especially for I.S.E. by Patricia H. Suter, Del Mar College.

Purpose

The purpose of this research was threefold. First, examination was made of the process of divergent thinking. Second, this research concentrated exclusively on the social behavior of the study group. And third, research concentrated on the intellectual activity of students in different social contexts.

Rationale

The assumptions underlying this work included the following:

--Students differ in divergent thinking potential.

--Grouping of students according to their scores on divergent thinking potential scores would have a significant effect on their ability to approach a scientific problem creatively.

--The verbal interaction (the amount and kind of verbal behavior of individuals) would differ when students were grouped homogenously or heterogeneously according to their scores on "divergent thinking potential."

Research Design and Procedures

One hundred six students from eight chemistry classes, involving four teachers, were the subjects of the study. The schools were

similar, being urban schools with predominantly middle-class students and good academic reputations. All students had studied chemistry for three-fourths of the school year and none had studied the topic which was the subject of the inquiry. Silent films were used to present the problem to be studied.

The students were classified on the basis divergent thinking potential into high, medium, and low potential. The students were told that they were in an experiment to see how they went about solving science problems. They were then divided into groups of four, shown a film, and asked to discuss, as a group, two questions:

1) As a group, list and discuss as many different explanations as you can for what you saw happening in the movie.

2) Choose one or more explanations and verbally describe what you would do in the laboratory in order to either test the explanation or get more data.

Both homogeneous and heterogeneous grouping was used and students were placed in both kinds of groups at different times. Six types of verbal interactions were correlated with eleven types of statements which served as the variables in this study.

Findings

Major findings reported are summarized below:

--Average number of statements made in each group was about 100.

--Cognitive-memory statements made up about three-fourths of the overall average.

--Divergent thinking potential accounted for about 10 percent of average number of statements. This was nearly the same percentage as cognitive thinking statements.

--In heterogeneous groups, students with high divergent thinking potential generally produce more interactions than students

0 0 8

with medium potential. Interactions in homogeneous groups are fewer, but difference is small.

--In homogeneous groups this trend is not followed.

--Students with medium divergent potential seemed to show more interactions in homogeneous groups.

--Results lead to the conclusion that grouping students homogeneously with respect to divergent thinking potential is useful in encouraging certain types of verbal interaction including divergent thinking.

Interpretations

The author emphasizes the importance of divergent thinking to scientific inquiry, especially because development of this approach is virtually ignored in today's schools. He also points out that divergent thinking is strongly related to creativity, transfer of learning, and problem solving.

ABSTRACTOR'S ANALYSIS

This article is an account of a study to relate students' scores on a divergent thinking potential scale with their verbal interactions in an inquiry or problem-solving situation. The study involved 106 students who actively participated in the project in that their responses to selected problems were observed and measured. This could be described as action research taking place within an existing school setting with the researcher controlling the problem selection and the classification of the students into groups of varying divergent thinking potential on the basis of prior tests and also controlling the categorizing of the responses of the students into some 11 verbal interaction areas. The student responses were balanced against six hypotheses and correlations tabulated.

The researcher found that students who scored high on tests of divergent thinking ability generally produce more verbal interactions than students with medium divergent thinking potential than students with low divergent thinking potential in heterogeneous groups. In homogeneous groups this expected trend is not followed. In the latter groups he seems to have found that these groups are advantageous for students with medium divergent thinking potential. The differences for the students with high divergent thinking potential are very small, which leads the researcher to conclude that "grouping students homogeneously with respect to divergent thinking potential is useful in encouraging certain types of valuable verbal interaction including divergent thinking."

The study seems to be one which involved a good deal of work on the part of the researcher, both in grouping the students with regard to divergent thinking potential and characterizing the extensive verbal interaction the groups had following the presentations of the problems the students were asked to consider. The researcher states that the study derives its importance from the fact that: (1) divergent thinking is virtually ignored in favor of verbal comprehension in today's schools, (2) divergent thinking is strongly related to creativity and transfer of learning which are of general interest, and (3) divergent thinking is strongly related to inquiry and problem solving which have become both established instructional strategies and desired educational objectives for the teaching of science in the elementary and secondary schools.

The researcher gives several references in his bibliography which detail his methods of grouping the students in the project into three groups (high, medium, low) on the basis of divergent thinking potential. Some of these are not readily available and are therefore not very useful (i.e., most universities do not have copies of doctoral dissertations unless they are the parent institution).

It has been impossible to date for this abstractor to obtain copies of the following references from four colleges and universities in this

area as well as the public schools: Aschner, M. J., Gallagher, J. J., Perry, J. M., Afsar, S. S., Jenne, W., and Farr, H. A System for Classifying Thought Processes in the Context of Classroom Verbal Interaction. Institute for Research on Exceptional Children, February 1965, University of Illinois, Cooperative Research Project No. 965, U.S. Office of Education Department of Housing and Economic Welfare; Torrance, E. P. Torrance Tests of Creative Thinking: Directions Manual and Scoring Guide: Verbal Test Booklet A. Princeton, N.J.: Ginn and Co., 1966, (a); Torrance, E. P. Torrance Tests of Creative Thinking: Norms-Technical Manual. Princeton, N.J.: Ginn and Co., 1966 (b).

It would have been productive if the author had included a brief summary of the tests he used for his classification.

Arlin, Marshall and Ian Westbury. "The Leveling Effect of Teacher Pacing on Science Content Mastery." Journal of Research in Science Teaching, 13(3): 213-219, 1976.

Descriptors--Chemistry; *Educational Research; *Instruction; *Pacing; *Programmed Instruction; Science Education; Secondary Education; *Secondary School Science; Teaching Methods

Expanded Abstract and Analysis Prepared Especially for I.S.E. by Elizabeth Kean, University of Wisconsin-Madison.

Purpose

This study explored the extent to which teachers' control of the pace of instruction influenced learning rates among high school students. The authors investigated a proposed leveling effect on learning rates brought about by teachers' pacing decisions, such a leveling effect being demonstrated by ablest students learning more efficiently when freed from teachers' control of the quantity of material to be learned, and the time allotted for learning. Thus, the study compared learning rates and learning rate variance under teacher-paced classroom instruction and under self-paced, individualized instruction.

The following specific hypotheses were proposed:

1. Rate variance would be significantly lower under teacher-paced instructional management than under self-paced instructional management.
2. The total group learning rate would be significantly lower under teacher-paced instructional management.
3. The leveling effect of teacher-paced instructional management would be minimal for slow students and most disadvantageous for potentially fast students.

Rationale

Managerial decisions that teachers make include the number of topics the class would "cover," selection of learning goals, the pace of instruction, who would get attention, concern, support, encouragement, etc. from

the teacher, when mastery was accomplished and the class could move onto the next topic, etc. Such decisions were assumed to be under the conscious (or unconscious) control of the teacher.

This study is related to previous work by Dahllof (1971), Wiley (1973), Anderson (1973), and Barr (1974) among others, which had related teacher management and deployment-of-instruction decisions and student achievement. Dahllof contended that teachers seek to impose a uniform, minimal level of achievement upon their classes. He has further suggested that to achieve this level, attention is directed to the less able students ("steering criterion group," somewhere between the 10th and 25th percentile on achievement scores) at the expense of the more able student. Lundgren's (1972) study was cited as confirming the existence of the "steering criterion group." The assumption appears to have been made that if teachers focus on the needs of low ability students, this will "certainly deflect teachers' attention, instruction, and press for achievement" away from more able students.

Barr's work (1974, 1975) in reading was cited as demonstrating that abler students in slower paced groups achieved significantly less than would be expected of them on the basis of their aptitude. A significantly greater variance in achievement was noted in classes where students were grouped by ability levels than in nonability-grouped classes. The greater variation in achievement in grouped classes was presumably due to the increased advances made by abler students.

This research was an attempt to extend the above work to science classes and specifically to look at the effect of teacher pacing. The authors assumed that they could rank order students by ability by ordering them according to learning efficiency (i.e., the more able students would learn more rapidly).

Research Design and Procedures

The subjects of the study were 68 students assigned to groups in two previously reported mastery learning experiments. Students were not randomly assigned, but the groups were similar in race and sex. The

instructional pattern for one group of 31 students was described as "teacher-paced (traditional)." These students were high school juniors enrolled in a chemistry class in an upper-middle-class suburban school district. The second group of students (37 high school sophomores enrolled in a biology class in a nearby, similar community) received "self-paced (traditional)" instruction.

The material studied by students consisted of seven chapters of an artificial, hierarchical pseudo-science in which principles and rules govern interrelationships of the elements of the system. No student would have previous knowledge of this pseudo-science, and thus, prior knowledge of material would not confound achievement results. Description of the material did not indicate whether the material in chapters was cumulative. In the self-paced format, students studied seven chapters of programmed material at their own pace. At the end of each chapter, students took a quiz on that chapter. Two teachers monitored the group, supervised quizzes, and walked about encouraging students to stay on the task. It was not reported whether teachers answered questions or performed any other typical teacher-type tasks.

In the teacher-paced treatment, students were taught the same seven chapters in a lecture-discussion format. Students were not given any written materials but were required to take notes on lectures as their source of information. When teachers determined that students were ready (criteria not given), the quiz was given to the class as a whole. Each chapter and quiz was given during the 50-minute class period of the chemistry class. The actual learning time was not reported, nor was the time interval between learning and testing, or the time interval between chapter presentations. No information about specific activities during the discussion periods was provided.

Each chapter quiz contained approximately 10 units of information. The "learning rate" was defined as the total number of concepts learned (answered correctly on the quizzes) per hour. An individual time was used for self-paced students; the teacher-controlled group time for the teacher-paced students. Odd-even reliability of quizzes (corrected by the Spearman-Brown formula) in measuring achievement and, by inference, the learning rate was .91. For analysis, students within each group

were divided into thirds (a high, middle and low group) on the basis of their learning rates

Findings

Hypothesis 1: Learning rate variance of the teacher-paced group (7.4) was significantly less than that of the self-paced group (75.9), $p < .01$.

Hypothesis 2: Group learning rate mean for the teacher-paced group (19.2 units/hr.) was significantly less than that of the self-paced group mean (25.0). An independent t-test showed the difference to be significant at the $p < .01$ level. The authors stated further that scores on science aptitude tests of general science reasoning ability and ability to read and comprehend science-related materials (previously given to these students) showed that the self-paced group had scored lower than the teacher-paced group.

Hypothesis 3: To determine if progress of abler students had been retarded, scores of all students were combined, and the sample divided into thirds on the basis on learning rates. Of the 20 highest students, 17 were in the self-paced group, whereas in the bottom third, only 7 were of the self-paced group. The X^2 value of 14.8 was significant beyond the .01 level, indicating the maximum detriment was to the faster students under teacher pacing. There was little effect on the learning rate of lower or middle third learning rate students.

Interpretations

The authors mentioned that results should be interpreted with caution because of nonrandom assignments, small sample size, short time period (7 days), and imaginary nature of the subject matter. However, despite these caveats, they concluded that the study offered support for the view that teacher pacing depresses mean achievement of students and narrows achievement variance, and that the decreases are at the expense of retrading progress of the ablest students. In summary, they state

that this study is an indication of the powerful influence of management factors such as pacing and steering criterion groups on the classroom.

ABSTRACTOR'S ANALYSIS

This is an example of process-product research. Two groups were assigned to different treatments, and the results were measured in terms of student performance on some achievement test. In this research, the outcome was "learning rate," the number of discrete items answered correctly by students per hour. The tests apparently were given to students immediately following the learning period. If this is true, then the tests measured short-term as opposed to long-term learning. Would the results have been the same if learning had been measured at different time intervals? Would the greater time spent on each chapter by the teacher-paced group translate into better retention over a longer time? We do not know.

Another feature of the learning outcome that was not well defined was the type of questions asked on the chapter tests. Were the questions recognition, recall, or both, of factual information? Were students expected to recall or use learned rules? Were there any questions that involved higher level skills such as extending material to new situations, problem-solving skills, etc.? Again, we do not know. The tenor of the article seemed to suggest attention to recognition/recall rather than more difficult or complex skills.

Both the duration of learning and the type of learning desired are important questions for science educators. What are the aims of teaching science? The management decisions that the authors discuss may be looked upon as basically curriculum decisions: what shall we teach and how shall we teach it, and to reach what goals and objectives? Studies such as the one reported here can provide information as to appropriate mechanisms for achieving specific limited objectives. However, one should be aware that the accomplishment of one objective may run counter to others. For example, in this study, short-term mastery of conceptual content would

seem to have been accomplished most efficiently for the most able students by self-paced instruction. Yet if a major objective of the science class were to increase students' facility to ask questions or to devise and defend a problem-solving strategy, then self-paced instruction (individual) might be less appropriate.

I would heartily endorse the cautions of the authors against translating the results of this study (or others like it) to classroom prescriptions. Many questions remain on a philosophical basis (to which classroom groups should attention be focused?) or a practical basis (how are "able" students or students most efficient in learning by themselves without teacher intervention to be identified?). Until such questions are answered, application of the results of this study cannot be made rationally.

Dahllof and others have introduced the terms management decisions, etc., as new constructs for us to consider. How do these constructs differ from what could be called simply curriculum decisions? Are we calling old ideas by new names? Does this serve any purpose?

Data Analysis

The authors used parametric statistics to analyze data on learning rates and learning rate variances. When discussing the mean group learning rates of self vs. teacher-paced classes, the authors bring in results of an undescribed science aptitude test which they claim showed the self-paced classes to have "lower attainment of seemingly relevant skills." This may indeed be true. However, in the absence of any further information, it is questionable whether this should have been mentioned in the "Results" section. We are unable to judge what the test might or might not have measured.

Treatment

If I were asked to replicate this piece of research, I could not do so because I would not know exactly what the two treatments were. The

authors offer us labels--teacher-paced and self-paced instruction--but do not expose for us the essential characteristics of such instruction. As Dubin and Taveggia (1968) found when they attempted to compare discussion vs lecture formats for college instruction: there is no consensus as to what exactly constitutes a discussion class, a lecture class, etc. We are left with a black box approach: do "something" to students, and see what happens to them. What within the black box causes the observed effects is a mystery.

Consider the complexity of the learning situation. In their efforts to set up a teacher-paced vs student-paced instruction, other differences occurred. The self-paced group was forced to gain information from a written presentation, while the other was required to process aurally from a teacher's presentation and to take notes on this material. From which type of presentation did the latter students actually learn? In the self-paced group, students needed to be active learners and could make decisions about when they were ready for the tests. In the teacher-paced group, how were the students kept on task? What method of presentation had the students been exposed to in previous science classes? Is a novelty effect (in method of presentation) working to make one type of instruction temporarily more efficient? On what basis did teachers decide that students were ready for testing? What was the nature of social interactions that were operating between students and teachers?

How was this study related to the study for which data had been initially gathered? Care should be exercised, when using data collected from one study for purposes of another, that biases are not carried over. For example, could teachers have been slower paced to ensure maximum mastery by their students for purposes of the other study?

Which of the above factors worked in favor of efficiency for the teacher-paced group or the self-paced group? Which of the above factors were responsible for the results of this study? Which other factors could have been involved? We simply do not know, both because no information is provided, and because there appears to be no way of separating out the variables that were operating.

The problem for authors of providing adequate information in journal articles is a real one. Editors are clearly saying "keep it short." At what point does conciseness conflict with completeness? Is the lack of information in this article due to enforced brevity or lack of adequate definition and control of variables? The reader cannot tell.

If the classroom situation is not well described, the results of studies such as this will always lack generalizability and applicability. However, even when the teaching situation is well described, the causal links between teacher actions and decisions and the characteristics of learning by students may prove elusive because of the complexity of the classroom situation.

REFERENCE

Dubin, R. and Taveggia, T. "The Teaching-Learning Paradox." Eugene, Oregon: Center for the Advanced Study of Educational Administration, 1968.

Combs, L. L. "A Teaching Approach to Physical Chemistry." Journal of Research in Science Teaching, 13(5): 467-472, 1976.

Descriptors--*Chemistry; Chemistry Instruction; *College Science; Educational Research; Higher Education; *Instruction; Motivation; Science Education; *Science Instruction; *Science Courses; *Student Motivation

Expanded abstract and analysis prepared especially for I.S.E. by J. Dudley Herron, Purdue University.

Purpose

The author compared the effectiveness of two patterns for managing instruction in a physical chemistry course.

Rationale

No reference is made to previous research and it is not clear that the author is aware of research that might illuminate the arena in which the research was conducted. The frame of reference appears to be the author's thoughtful consideration of possible consequences of various classroom procedures. The author makes statements such as:

> "Some ... students require external stimulation in terms of teacher-set goals to arouse their interest.... Other students will rebel at such a regimen...."

> "...Is it possible that allowing drop grades is conductive to a more lax attitude on the part of some students and actually detrimental to proper motivation?"

> "...If lectures were taped, would more students miss class and listen to the taped lectures, or would they come to class and listen more and write less during class?"

Research Design and Procedure

On the basis of rhetorical questions such as those mentioned above, the author compared the performance of students during the first semester of a two-semester course with the performance of the same students during the second semester of the two-course sequence.

The following descriptors were applied to the first and second semester courses.

First Semester	Second Semester
a. teacher pacing	a. self-pacing (optional)
b. dropping of grades	b. no dropping of grades
c. pop tests	c. no pop tests
d. infrequent testing	d. frequent testing
e. no taped lectures	e. taped lectures
f. complete course outline given to students	f. partial outline furnished for students to complete

During the first semester, four hour-exams were given and one was dropped. Seven homework sets were assigned and ten pop tests were given. Lectures were not taped, no self-pacing was allowed, and complete outlines were given to the students.

During the second semester, "a complete objective outline of the course was given, but [students] were required to prepare a more detailed study outline. Eight tests were scheduled, and an early test on each unit was allowed if the homework problems ... were first completed.... If an early test on a unit was taken, a grade of 90 was required to consider the unit complete. If such a grade were not made, another test on that unit would have to be taken at the scheduled time." Lectures were taped and taped previews and problem-solving aids were also available. Students could pace themselves but lectures were regularly scheduled and few students missed them.

There was no research design. An attempt was made to compare the average grades obtained each semester.

Findings

The author reports means, standard deviations, and histograms for scores on each exam given during the first and second semester of the course. During the first semester, the mean score on the four exams ranged from a low of 52 percent to a high of 76 percent and there was no evidence of improved level of performance during the course of the semester. During the second semester, the mean score on the eight exams ranged from a low of 71 percent to a high of 90 percent and, after the first four exams, the author interprets the results to show a steady improvement in performance.

Interpretations

The following quote seems to summarize the author's evaluation of the experience.

> "This multiformity approach [the author's term for the organization used during the second semester of the study] has been utilized now for five semesters, and its value is being repeatedly revealed. It requires more individual work than the traditional approach and quickly removes students from the course who are not willing to work. ...[C]lass attendance is better under this system than the traditional method... [I]t is imperative that [students] are taught to use the taped lectures properly...A few students will self-pace one or two of the units...however, the students seem to need the lectures and feel more comfortable with instructor-set goals...[F]requent testing and no dropping of grades [motivates students to keep up] much better [than pop quizzes.].... Since every student has different weaknesses and strengths, their prepared outlines are much more useful to them than a general outline.
>
> "This multiformity approach...combines the best aspects of the self-paced course designs and the formal lecture course designs

to present a comfortable environment for any serious student" (p. 472).

ABSTRACTOR'S ANALYSIS

My overall impression is that a conscientious professor of chemistry wanted to improve his course in physical chemistry. He gave serious thought to the possible reaction of students to various modifications that he might make in his course. On the basis of that reflection, he instituted various changes that appeared to be sensible. In an effort to assess the wisdom of his decisions, he observed--as best he could in a situation that precluded true experimentation--the students' performance before and after the decision to change the organization of the course. The professor then made some conjectures--based, no doubt, on informal observations as well as formal ones--concerning the value of the course modifications.

There is a gnawing temptation to castigate the author of this study for numerous violations of standard behavioral research methodology. Absolutely no research has been cited, in spite of the fact that there has been a great deal of research on instructional variables such as those manipulated in this study.* There was no research design to control for the various threats to internal and external validity of the study. [Although the tests given before and after the modifications in course organization provide an aura of a time series study, the fact that each observation (test) was different and of unknown difficulty makes it impossible to interpret the study under the time series paradigm. Nor does the author claim that such an interpretation is possible.]

By making several modifications in course structure simultaneously, it is impossible to know which modification (or combination of modifications) had an effect--assuming that it is possible to demonstrate some effect.

---

*The Summary of Research in Science Education published by John Wiley and Sons as a supplement to Science Education over the past several years provides numerous examples of such research.

In reporting the results, the author describes "a comparison of performances (2) by the propagated error formula" as though this is a standard technique that will be recognized by any alert reader. Apparently it is not. After consulting with several colleagues in educational research, in statistics, and in chemistry, I was unable to learn what the author had done. (The notion of propagation of error is a familiar one, but what the author has done is not at all clear from the report.)

All of the above are obvious problems with this research report. At least some of the problems should have been corrected by the referee who recommended acceptance for publication. Still, there are more important points to consider.

We are in the business of improving science instruction. Most of us believe that improvement is most likely to come when we carefully scrutinize what we are doing--that we make rational judgments based on data, rather than whim. In other words, we want to be "scientific." For that reason we have borrowed research designs and the methods of statistical inference from other fields and applied them to education. Unfortunately, these methods presume an ability to manipulate variables and to randomize treatments that often extends beyond possibility. We are often faced with the choice of collecting data which are subject to reasonable doubt, or making decisions on the basis of no data at all. Furthermore, if we share our considered judgments based on limited data and inadequate research design, we expose ourselves to "the gnawing temptation to castigate" rising in the professional souls of our colleagues.

Carver (1978), in presenting the case <u>against</u> statistical significance testing, makes some cogent points that seem pertinent in reviewing this article.

> "Some people avoid classroom research, for example, because the use of intact groups usually violates fundamental assumptions of analyses of variance; others go ahead and calculate <u>p</u> values

from these analyses of variance even when they are completely erroneous. It would be better to disregard statistical significance in these situations where the classroom is the unit of analysis and where there are too few classrooms to get statistical significance" (Carver, p. 394).

While recognizing the very serious limitations of this study and while recognizing the many competing hypotheses that might explain the observations that are reported, I am not ready to ignore it. The author's explanation of the observations is at least as good as any other. The tenor of the report suggests that the author was seriously attempting to make objective judgments rather than to promote a new, saving elixir.

If somebody can design an experiment to investigate the effect of the instructional system introduced in this study, I encourage them to do so. In the meantime, I would have no qualms about encouraging other professors of science to develop a course structure similar to the one instituted by the author of this study. Neither would I discourage them from attempting to collect data that might help them decide whether the changes are justified, even though the data are not based on the ideal of a true experiment. I could even encourage them to share their results, so long as they make no extravagant claims or exude a confidence unjustified by tenuous data.

REFERENCE

Carver, R. "The Case Against Statistical Significance Testing." Harvard Educational Review, 48: 378-399, 1978.

Eccles, P. and G. J. A. Deleeuw. "Information Processing and Pupil Achievement: A Model and A Study." Journal of Research in Science Teaching, 13(5): 383-395, 1976.

Descriptors--*Educational Research; *Elementary School Science; Elementary Education; Grade 6; *Instruction; Interaction Process Analysis; *Science Education; *Teaching Models; *Teaching

Expanded abstract and analysis prepared especially for I.S.E. by Willis H. Horak, University of Arizona.

Purpose

This study was designed to identify specific teaching behaviors which might be related to differences in pupil achievement. It was specifically limited to analysis of lessons in which the acquisition of new information and the processing of that information takes place. The major objectives of the lessons analyzed were therefore not on recall or consolidation of facts and principles but, instead, on students' acquisition of information for the purpose of developing new concepts and rules during the lesson period.

Rationale

A major proportion of this article was devoted to an explanation of the contextual model within which the study was designed. The model is concerned mainly with pupil-subject matter interaction under the direction of the teacher. It considers four processes of teaching: (1) providing fact/experiences, (2) describing, (3) relating, and (4) valuing.

Providing fact/experiences may be described as those subject matter interactions where an opportunity to perceive instances of concepts and relationships or to receive reports of such perceptions is provided. These experiences may be grouped according to the degree with which they may be influenced by teacher mediation and range from direct observation to verbal directions and discussions. Describing refers to verbalizing or otherwise communicating information when doing so does not require

the development of new concepts or an understanding of relationships not immediately recognized. Relating is defined to include such processes as explaining, defining, abstracting, generalizing, and others. This processing of the subject matter information is done by the teacher and pupils when the objectives of instruction specify learning of new concepts and principles. Valuing is defined as arriving at an independent judgment or exhibiting the independent capacity to perceive usefulness of concepts and generalizations in new situations.

There are three main types of school lessons described in this paper: (1) recitation lesson, (2) consolidation lesson, and (3) developmental lessons. Of these three types of school lessons the processes previously described apply only to the developmental lesson.

Research Design and Procedures

For this study 12 teachers taught a developmental lesson on four different occasions to groups of 10 sixth-grade children. The lessons were concerned with the concepts and principles of flotation. Transcripts of the lessons were organized into nine-second intervals and analyzed according to a system developed by Smith et al. (1967). In this system ventures and moves are analyzed. A venture is defined as a unit of discourse consisting of utterances dealing with a single topic and having one overarching objective. Causal, conceptual, particular, and procedural ventures were found in this series of lessons. A move is defined as the verbal manipulation of the content of instruction. The moves have been assigned to levels derived from the general model of the teaching process in science and social studies.

Since the instrument used is one which requires a lengthy training period to master, all lessons were analyzed by the investigator. Reliability coefficients were calculated on the scorings according to procedures described by Myer (1966). These reliability coefficients were all high and could be interpreted as a measure of the correlation between the mean scores for two sets of a randomly-sampled item administered at different times to the same sample of subjects.

The criterion measure for the study was pupil achievement. An overall Teacher's Class Mastery Score was calculated for each lesson. This score was defined as the proportion of those pupils not understanding at the beginning of the lesson who were judged as understanding the principle at the end of the lesson. There were two response papers for each child. These were examined by two judges who independently decided on the totality of the evidence on the paper if the child did or did not understand the principle.

Findings

The data for this study were analyzed according to correlational techniques. Of the 15 correlation coefficients calculated, three were significant. The analysis revealed a significant correlation between the number of higher-cognitive level information-processing moves and pupil achievement. A similar related significant result was obtained with pupil achievement when the time spent in higher-cognitive level information-processing moves was used as the unit of measurement. Lastly there was also a significant correlation between the number of information processing moves at Level 8 and pupil achievement.

Interpretations

First of all, the authors caution that, although the essence of teaching is to cause change in pupils, the design of the study did not allow them to formulate causal hypotheses. The correlational data do not imply cause and effect and in fact the work of teachers varies in a large number of dimensions and any change in one must be integrated into an overall pattern. The authors still felt, however, that it was a disappointment not to be able to demonstrate a positive relationship between pupil achievement and direct pupil experience. This is a type of activity so dear to many good science teachers. Possible explanations for the results of this study are the small number of teachers involved and the time limit. These may possibly affect the relevance

of different teaching behaviors. Still the amount of time given to higher-level thinking processes by the highest-scoring teacher still left adequate time for including direct experience as part of the lesson.

Lastly Rosenshine and Furst (1971) have stated that, for good teaching, classroom activities should be on higher cognitive levels, such as hypothesizing, evaluating, generalizing, synthesizing, comparing and contrasting. This study supports this "should" with a modest bit of empirical evidence.

ABSTRACTOR'S ANALYSIS

This reported study is most worthwhile as part of the initial studies necessary for model development in education. A number of educators and researchers (Reigeluth, 1978; Long, Okey, and Yeany, 1978; Hyman, 1970; Goldstein and Howe, 1978) have previously demonstrated or expressed beliefs that teaching can and should be viewed as a relationship among teachers, pupils, and subject matters. This paper and study provide one model for analyzing a number of hypothesized relationships among the variables. However, it is not the only model that may be useful and others, such as that proposed by Riegeluth, may also be worthwhile.

Overall, the article is very well written. The tables that are presented are most useful in helping readers interpret the relevancy of the described results. The ranges, means, and standard deviations provide the basis for a more in-depth post-hoc analysis of the study.

The validity of the study is also enhanced by the random assignment of students to teachers. Additionally, the unit of analysis being the teacher adds to the credence one may put in the results. These are two areas where other research studies do not always meet the conditions of the statistical research model employed.

The authors themselves discuss many of the problems related with this type of research. They also adequately explain procedures they employed to alleviate or minimize many of the problems. First of all, one notes that a single investigator analyzed and scored all the tape-scripts for the study. Thus one does not know for sure the validity of this investigator's interpretation of Smith's instrument. This is pointed out in the paper. The reliabilities reported for systematic bias in interpretation do show consistency of scoring. However, this does not address the validity question.

The small sample size may also cause interpretation problems. Again the authors point this out in their discussion. With this type of study, however, a larger sample size leads to a much larger additional input of time. One must carefully weigh this when designing such a research study. Replication of this study by another researcher will help alleviate both of these problems.

One last problem, not discussed by the authors, is the fact that the report does not adequately describe the scoring of the criterion measure. This alone makes replication of the study very difficult. Exactly what types of "totality of evidence" suggested understanding of the principles of buoyancy? If there was a consistent criterion used, it should be possible to generally describe this criterion.

This study adds significantly to the development and understanding of an educational model for analyzing science teaching. The fact that low cognitive level, direct type experiences had a low or a negative correlation with student achievement may be a "disappointment" to many science teachers and science educators. As is often the case when a study does not agree with pre-conceived ideals of teaching, one is inclined to criticize the research methodology. Such may be the case in this study. However, as pointed out, the authors are aware of many of the deficiencies of their study. These are discussed in depth and, in my opinion, are adequately taken into account in the interpretation of the data. It was not the purpose of this paper or study to determine cause and effect relationships between teacher behaviors and

student achievement. Instead it was the intent to describe a relevant information processing model and complete an initial investigation within the described model's framework. The authors are to be commended for succeeding in this endeavor.

From this study and the reported tables one may formulate many other research questions that may be investigated. We know that more time is necessary when students are working in an inquiry mode that involves discovery. Therefore time may be negatively correlated to student achievement on a short time basis. What about retention? An achievement test administered three or four weeks after the study may be quite informative. Such an achievement test would also be easy to administer. We may find that the reported results are quite misleading for long-term retention.

Also, what about transfer? Is there still a high positive correlation between high-cognitive level information processing moves of the teacher and student achievement on transfer items? These questions need to be researched to further outline the usefulness of the model.

A model of the type described should be useful for identifying, conceptualizing, and studying the ways in which teachers, students, and subject matters interact. There is research evidence that teacher types may be identified which are differentially effective with students. Cunningham (1975) has identified four such teacher types. Ebmeier and Good (1979) have also reported a differential effect on achievement between teacher types and student types. Thus the model may have to be modified to include such interactions. The teacher has attitudes, beliefs, and values that influence instruction. If teachers value certain methods, they change their teaching styles to accommodate these methods of instruction. If they do not value certain methods, they may use the instructional method but not in a consistent manner. Similarly, students value certain ways of learning. Overall means computed in research studies may mask many individual differences in students' ways of learning that influence achievement. For subject matter variables, criterion measures need to be conceptualized along continua basic to

the total science curricula. Through analysis of possible interactions within a model of instruction, a general theory may possibly be developed that considers individual differences of these three categories of variables as basic components of effective, efficient and appealing schooling.

REFERENCES

Cunningham, W. G. "The Impact of Student-Teacher Pairing on Teacher Effectiveness." American Educational Research Journal, 12(2): 169-189, 1975.

Ebmeier, H. G. and T. L. Good. "The Effects of Instructing Teachers About Good Teaching on the Mathematics Achievement of Fourth Grade Students." American Educational Research Journal, 16(1): 1-16, 1979.

Goldstein, M. and A. C. Howe. "The Use of Concrete Methods in Secondary Chemistry Instruction." Journal of Research in Science Teaching, 15(5): 361-366, 1978.

Hyman, R. T. Ways of Teaching. Toronto: J. B. Lippincott Company, 1970.

Long, J. C.; J. R. Okey; and R. H. Yeany. "The Effect of Diagnosis with Teacher or Student Directed Remediation on Science Achievement and Attitudes." Journal of Research in Science Teaching, 15(6): 505-512, 1978.

Myer, J. K. Fundamentals of Experimental Design. Boston: Allyn and Bacon Inc., 1966.

Smith, B. O.; M. Meux; J. Coombs; G. Nuthall; and R. Precious. "A Study of the Strategies of Teaching." Urbana, Illinois: Bureau of Educational Research, College of Education, University of Illinois, 1967.

Reigeluth, C. M. "A Methodology for Investigating the Interactions of Individual Differences and Subject Matter Characteristics with Instructional Methods." Report No. 67. Courseware Instructional Research Reports, Brigham Young University, 1978.

Lawrenz, Frances. "The Stability of Student Perception of the Classroom Learning Environment." Journal of Research in Science Teaching, 14(1): 77-81, 1977.

Descriptors--*Classroom Environment; *Educational Research; Learning Activities; *Measurement Instruments; Science Education; Secondary Education; *Secondary School Students; *Student Opinion

Expanded abstract and analysis prepared especially for I.S.E. by Chris A. Pouler, Hyattsville, Md.

Purpose

The Learning Environment Inventory (LEI) was designed to measure students' perceptions of their classroom environments (Anderson, 1973). Based on the assumption that learning environment remains stable throughout the academic year, researchers have considered a single administration of the LEI as "reliable, valid, and sensitive for assessing differences among or within classes" (p.77). The purpose of this study was to examine the validity of this assumption by measuring the LEI scores of students from various classroom environments during an academic year. Specifically three questions were investigated.

1. Are class mean scores for the LEI stable over time?
2. Are individual student scores for the LEI stable over time?
3. Do individual students who take the LEI more than once have different perceptions of the classroom environment than their classmates who take it only once? (pp. 77-8)

Rationale

As an important aspect of the learning process, the classroom environment must be accurately assessed if worthwhile data are to be discovered. The LEI has provided a simple and useful assessment of the classroom environment as perceived by students. However, since researchers tend to administer the LEI once during the school year, the assumption has

prevailed that students' perceptions remain constant. This study sought to question the assumption and, thus, strengthen or weaken findings and generalizations from previous studies.

Research Design and Procedure

Specifically the LEI determines those aspects of classroom atmosphere pertaining to management and student/teacher interactions. This study utilized a "modified version of the LEI containing ten scales related to classroom social situations" (p. 78). Each scale involved scoring items using numbers one to four which corresponded to the range of the continuum from strongly agree to strongly disagree. The scales measured each of the following variables.

1. Diversity: the extent to which the class provides for a diversity of student interests and activities.
2. Formality: the extent to which behavior within the class is guided by formal rules.
3. Friction: the extent to which conflict exists among the students in the class.
4. Goal Direction: the extent to which the goals of the class are recognized by its members.
5. Favoritism: the extent to which differential treatment of students exists in the classroom.
6. Difficulty: the extent to which the work of the class is perceived as difficult.
7. Democratic: the extent to which all students participate in class decisions.
8. Cliqueness: the extent to which cliques are present in the classroom.
9. Satisfaction: the extent to which students are satisfied with the class.

10. Disorganization: the extent to which the class is perceived as disorganized (p. 78).

The subjects represented students from physics (n=7), chemistry (n=5), biology (n=8), and nonscience (n=7) classes of western and central New York. The results were gathered during the 1974-75 school year from tests administered during the months of November, February, and April. To obtain valid results the LEI was given to a random-half of each class on each test date. Consequently the students who had taken the test one, two, and three times comprised the three sample groups. To test the null hypotheses for each of the research questions, MANOVA statistics were employed.

Findings

1. The hypothesis that class mean scores were stable over time was accepted. Further, since no significant interaction existed, type of course and stability were considered independent. There were significant differences due to the type of course, which the researcher expected.
2. The hypothesis that individual student scores are stable over time was substantiated. Again there was no significant interaction between type of course and stability. Significant differences existed for type of course.
3. The hypothesis that students who take the LEI more than once do not differ from students who only take the test once was proven correct. Further, no difference due to test sensitization was shown. Likewise, there were significant differences for the type of course.

Interpretation

Since students' perceptions of their classroom environments are stable, the importance of the learning environment has been illustrated. Previous

findings have also been reinforced. Future research can continue by quantifying particularly successful learning environments with teacher and course variables.

ABSTRACTOR'S ANALYSIS

The Learning Environment Inventory appears to provide useful data for researcher, guidance counselors, and classroom teachers. In a previous study Lawrenz (1976) used this instrument to compare the mean scores for biology, chemistry, and physics classes. For each scale of the instrument, significant differences occurred between at least two and often between all of the pair-wise comparisons. For guidance purposes these findings can be useful to match the student to an elective science course. Further, teachers with the knowledge of students' course perceptions can better develop appropriate course objectives and lessons. In the present study, Lawrenz continued her work by questioning the foundation of the original conclusions. The assumptions that students' perceptions as measured by the LEI remain stable was proven valid. Further, the pattern that perceptions vary with the type of course was repeated. Unfortunately the extent of the pair-wise differences was neither discussed or presented adequately.

The procedures utilized were sound. While the random half approach for each test administration insured good testing conditions, equal group size was lacking. Random placement of students into one of three groups (one, two, or three test administrations) would have provided three equal sized groups. Further, a more in-depth description of the population would have been useful. For example, were all biology classes exposed to a comparable text and teaching approach? Students taught traditionally should perceive their classroom environment differently than those taught via inquiry. (Or should they?)

Significant differences due to type of class were reported--as expected--but without further comment. A comparison of these results with those from the 1971 study would have been interesting and insightful. While

differences due to course type were not the objective of this study, the repetitive pattern of the finding does reinforce the previous generalizations made by Lawrenz. For the future, the differences of perceptions described by students of physics, chemistry, and biology may be as important as the stability of the classroom environment.

As for the LEI, classroom management techniques, nature of the course content, and interactions between students and teachers are measured. Future studies might attempt to further quantify the scales of the LEI. In addition, researchers might wish to correlate the reactions of various age groups and interests to each scale.

Especially important is the generalization made by the researcher that teacher or course variables could be manipulated to create a specific environment. Practically speaking, successful traits could be identified and passed on for pre-service teachers. Further in-service programs involving specific curricula would be valuable to teachers if students' perceptions were disclosed. Unfortunately, however, the LEI does not delve into students' feelings about a specific course. Perhaps the diversity of biology is constant throughout the year, but so what! Does this finding imply that diversity is good? With a stable learning environment does a student learn process as well as content? Further, as semester courses and modules increase in popularity, how will the stability pattern be affected? Similarly, is stability constant for students who are exposed to the range of teaching methods from individualization to lectures throughout the year? A classroom may provide a stable environment but does this imply students learn best as a result? Questions abound regarding each scale of the LEI and the corresponding relationship to learning.

In summary, this study is worthwhile since the usefulness of the Learning Environment Inventory has been established. Future studies should determine how stability of learning environment relates to the most crucial outcome of education--learning.

REFERENCES

Anderson, G. J. The Assessment of Learning Environments: A Manual for the Learning Environment Inventory and the My Class Inventory. Halifax, Nova Scotia, Canada: Atlantic Institute of Education, 1973.

Lawrenz, F. P. "Student Perception of the Classroom Learning Environment in Biology, Chemistry, and Physics Courses." Journal of Research in Science Teaching, 13(4): 315-23, 1976.

Lloyd, Gerald. "An Experiment with the Teaching of Science to Nonstreamed Classes." The School Science Review, 58: 355-360, 1976.
Descriptors--*Course Evaluation; Educational Research; *Heterogeneous Grouping; *Individualized Instruction; Instructional Improvement; *Intermediate Grades; Scientific Attitude; Secondary Education; *Secondary School Science; *Student Attitudes

Expanded Abstract and Analysis Prepared Especially for I.S.E. by Dorothy Gabel, Indiana University.

Purpose

The author's purpose for conducting this study was to determine whether the use of a variety of worksheets for teaching science in a self-paced setting was more effective than using identical worksheets in a group-paced situation. Worksheets for the self-paced classroom were devised for different ability ranges, whereas worksheets for the group-paced class were the same for all students.

Rationale

In heterogeneous (nonstreamed) classrooms, one of the prime aims of teachers is to encourage all children to work at capacity. This requires that instruction be individualized to some extent. One method of individualizing instruction is to supply students with worksheets (or programmed texts) that students complete at their own rate. If identical worksheets are provided for all students, however, they may not be sufficiently challenging for the better students or may be too difficult for the slower students. This study examined the effectiveness of using worksheets devised for different ability groups. No references to previous research in this area were given in the report.

Research Design and Procedure

The sample consisted of eleven- and twelve-year-old students in one school located on the outskirts of Bristol. Students in the school were divided into groups of 30. Two of these groups were selected for

the experiment, one constituting the experimental group and the other the control. Students in the experimental groups were divided into three subgroups according to intellectual ability.

In the experimental classroom, students were assigned to learn science through the use of worksheets that were prepared for the low, middle, and high ability subgroups. Worksheets varied according to: 1) The reiteration of experience in skills and knowledge that had been covered in previous work, 2) the depth of reasoning and discussion expected within the subgroups, 3) the number of difficulty of problems posed for immediate solution, 4) the complexity of the written responses required, and 5) the practical skill needed for the necessary experimentation.

Students in the control classroom were taught as closely as possible to the recommendations of the Nuffield Guide. This was generally class-paced instruction with the same worksheets for all children in the room.

Three instruments were administered to the students. Achievement was measured by a self-produced test that contained items requiring knowledge, comprehension, and higher mental processes. This test had a reliability of .82 to .91. It was administered immediately after the one-term course, and again six weeks later. An attitude measure and a mental ability test were also administered but no information about them is given in the report.

Findings

On the achievement test, results indicated that students who had studied using the individualized material achieved significantly higher ($p < .05$) than children being taught using the conventional approach. Analyses of the subsections of the test (knowledge, comprehension, and higher mental processes) showed that this was also true for the knowledge and comprehension subsections. The above results were consistent for both the immediate and delayed posttests.

In the affective area, there were no significant differences on gain score means between students in the two groups on their attitudes toward science in life, science in society, learning activities, or processes of science. The only significant difference was in their attitude toward school where the group-paced class became more positive and the self-paced class more negative.

In addition to the above instruments, opinions of students and teachers were included in the report. On the whole, students did not favor the individualized approach. The teacher who used this approach also found it less desirable than a more conventional approach.

Interpretation

Results of this study indicate that students' achievement in science, particularly on the knowledge and comprehension level, is enhanced if they use worksheets geared to their mental ability in a self-paced classroom. This, however, is not the only consideration that must be made in teaching science. The worksheets did not enhance the student's overall attitudes and, in fact, they became more negative toward school. This finding, coupled with the negative responses that students gave on the opinionnaire about the method, suggests that the use of worksheets on a continual basis (for a term) is not desirable. Using individualized worksheets for shorter time intervals and coupled with other diverse teaching strategies may be a more satisfactory method of teaching science to eleven- and twelve-year-old students.

ABSTRACTOR'S ANALYSIS

This study should be considered a pilot study that gives some indication that the use of worksheets prepared for different ability groups increases achievement in nonstreamed classrooms. The major reasons for considering this a pilot study are that the sample is very small (approximately 30 students per classroom in two classrooms), all students were from the same school, and each had a different teacher. The 30 students

were subdivided according to ability giving only 10 students per group. This is a small sample size. The fact that only two teachers participated in the study makes it difficult to generalize to other teachers. The results may be due to the interaction of the teachers with strategies. The study needs to be replicated using a broader sample of children and with more teachers participating.

This is not to say that the study is without merit. Much energy went into the preparation of the three forms of student worksheets for an entire term. The fact that the experiment extended over a long period of time is also commendable as this unfortunately is not a common practice. Noteworthy also is the fact that retention was measured six weeks after the completion of the term.

In addition to broadening the sample of children and involving more teachers, this study and the report could have been improved in the following ways:

1. Instrumentation. Specific details about the instrument administered should be included in the report. In this report the name and reference for the mental ability and the attitude tests were not given. The reliabilities should also have been stated, as well as the formula used to calculate the reliabilities. The number of test items is also helpful.

   If a test is analyzed in subsections, the reliabilities of each subsection should be examined. The entire test may have a high reliability but particular subsections might be below an acceptable level. This would apply to both the achievement test and the attitude test in this study.

2. Analyses of data. The data in this study are presented in terms of percentages with no information given about the number of items on the test. The latter information is of interest to the reader who might wonder if a 50 percent score meant getting 1 out of 2 right or 50 out of 100. If the test is primarily multiple-choice items as is indicated in the report, one questions how much was learned by students in the low-VRQ group control group when their scores were

around 20 percent (at the chance level). Because n is not given, one also wonders about the size of each of these groups. Were there 10 students in each or were the average groups larger and above and below average groups smaller?

Data for the attitude test were analyzed according to mean gain-scores. Analysis of covariance would have been a more appropriate method of analysis.

3. Interpretation of the data. If data appear to be inconsistent with the general results, an explanation is generally called for. In this study the mean scores in the test taken six weeks after the other test appear to be inconsistent across groups. Generally, the retest scores are about the same as for the first test. One wonders, however, why the scores of students in the control groups would increase substantially in four out of nine cases whereas this was true for only two of nine cases for the experimental group.

Mayfield, J. M. "Factors Affecting Rationality in the Discussions of a Problem by Small Groups of Secondary School Students." Science Education, 60(2): 173-183, 1976.

Descriptors--*Educational Research; *Group Discussion; *Group Dynamics; Instruction; Science Activities; Science Education; Secondary Education; *Secondary School Science; *Teaching Methods

Martin, J. F. "Analysis of John M. Mayfield's Factors Affecting Rationality in the Discussion of a Problem by Small Groups of Secondary Students." Science Education, 60(2): 185-192, 1976.

Descriptors--*Educational Research; *Group Discussion; Instruction; Problem Solving; Science Education; Secondary Education; *Secondary School Science; *Small Group Instruction; *Teaching Methods

Expanded abstract and analysis prepared especially for I.S.E. by Dorothy L. Gabel, Indiana University.

These two articles are considered together because one is a critical review of the other. The analysis by Martin will be incorporated in the *Abstractor's Analysis* because of the nature of the article.

Purpose

In conducting this research Mayfield had as his primary purpose the exploration of factors operating within small groups of adolescents that affect the rationality of the groups as they attempted to solve a particular problem.

Rationale

Although small group instruction is frequently utilized in the secondary schools, there is little theoretical basis justifying it as a suitable mode of instruction. Studies by Walberg and Ahlgren (1970) and Walberg and Anderson (1968) have produced post hoc descriptions of correlations between learning outcomes and the classroom social climate in secondary physics classes, but direct observation was not a part of that research. A study by Hurd and Rowe (1966) indicated that in high school biology,

some incompatible groups reduced intragroup friction by spending much of their time on social-emotional behavior whereas others avoided intragroup tensions by concentrating on their task.

This study closely examines intragroup behavior using Bales' spatial model of group structure (1950). Utilization of this model involves mapping the behavior of each group member in a three-dimensional "group space," the axes of which are behaviors: Forward-Backward (task-oriented versus non-task-oriented), Up-Down (dominant versus submissive), and Positive-Negative (friendly versus unfriendly).

Research Design and Procedure

This exploratory study used a naturalistic, observational approach in which 60 eleventh and twelfth grade students from a public coeducational high school and 25 students from a parochial girl's school were divided into groups of five students each. Each group was given the same problem to solve: the NASA "moon survival" problem by Hall (1971). The problem each group was to solve was the listing, in sequence of importance, 15 items necessary to survive on the moon in an emergency landing.

One day prior to the group solution of the problem, students completed a 15 item multiple choice pretest of relevant information about the problem solution and were asked to solve the problem individually. On the day of the group solution, students completed Bales' Interpersonal Rating Form A (revised) after solving the problem. On the following day, they completed a posttest that was identical with the pretest, and also solved the NASA problem individually once again.

Findings

A. Problem Solving

1. Most individual's initial solutions were not as correct as "their group's" solution.

2. Most individual's second solutions tended toward the group's solution, even when the group's solution was less correct than the individual's initial solution.

B. Group Structure

Three types of groups were identified.

1. Type I: five groups - low scatter, clustered interpersonal ratings.
2. Type II: three groups - intermediate scatter and clustering of interpersonal ratings.
3. Type III: nine groups - wide scatter, no clustering of interpersonal ratings.

C. Interaction Effect

1. Type I
   Tended to have rational discussion. In four out of five groups, group solution was superior to that of best member.
2. Type II
   One of three groups had group solution superior to that of best member, the remaining two being equal.
3. Type III
   Tended to argue on nonrational grounds. Six out of nine groups gave solutions inferior to that of best member.

Interpretations

Generalizations from this study must be made with caution because students were not randomly selected from the general population. Some interpretations about the inner-workings of the groups were that the nonrational grounds on which students argued could be classified as concessions, force, rules-of-thumb and emotion. Rational discussions were useful in two ways: individuals who made rational suggestions facilitated their own ascendancy within the group, and the group as a whole developed a task-oriented pattern. The phenomenon of devaluation that occurred in every group occurred

when group tensions were high "due to loss of problem mastery of ascendant group members." Once devaluation took place, the social value to the individual of having his ideas accepted diminished and ascendant group members shared social status rewards with those of lower status. Rationality of discussion generally declined.

The findings indicate that the idea of small-group instruction does not necessarily lead to more verbal participation in learning by all students, nor does it improve social development if some group members follow submissively. Students need to be made aware of social interactions that do occur and their effect on learning. Pitfalls can then be overcome to make small group instruction more valuable for all.

ABSTRACTOR'S ANALYSIS

The major strength of this study is in the effort the researcher made to examine in detail the social structure of groups in a problem solving situation. Too frequently only the overall effects of research are reported and little is known about the inner workings or interactions among individuals comprising groups. Of great merit also is the author's precautions in generalizing beyond this group because of the nonrandom nature of the selection of the sample.

The following questions about the research are raised by both Martin and this reviewer.

1. It is not clear how either the rationality of the approach or the ease with which members worked together were affected by the pattern of asking students to complete the moon survival problem individually before solving it as a group. Any prior commitment a student would have toward his/her individual solution might affect the rationality of the group.

2. The NASA problem is not representative of the types of problems a student frequently encounters in science and students may react differently to its solution. In other words, results may not be generalizable to other science problems.

3. The interactions of the group may be affected by "group history." In this particular case the students were random collections of preacquainted individuals who came together for one short problem-solving session. How these groups would react to a longer time together solving a variety of problems as would ordinarily be encountered in a regular classroom is unknown. Generalizations from this one experiment cannot be extended to the normal classroom situation.

4. Group members used their own perceptions of how they reacted in the group and these may be lead to inflated or deflated scores. Reports by observers of the group while it was in action may give unbiased ratings for members in the group.

5. Bales' group space model may not be applicable to adolescent groups. Bales' model (as noted by Martin) has been developed using college students and adults and assumes a degree of personal stability. Adolescents may not have this required stability.

In summary, Mayfield has contributed to the area of science education in carefully analyzing the processes that occur in small group instruction. Martin has added to this contribution by his careful analytical critique of Mayfield's work. Not only did he analyze the article, but he used Mayfield's full thesis in preparing his analysis. Additional research studies are needed with more randomly selected samples not only to determine the applicability of Bales' group space model to adolescents but also to determine the effects of small group instruction on achievement and social development.

REFERENCES

Bales, R. F. "Interaction Process Analysis: A Method For the Study of Small Groups," Reading, Mass.: Addison-Wesley, 1950.

Bales, R. F. "Personality and Interpersonal Behavior," New York: Holt, Rinehart and Winston, pp. 7-10, 29, 1970.

Hall, J. "Decisions, Decisions, Decisions," Psychology Today, November, 1971.

Hurd, P. de Hart and M. B. Rowe. "A Study of Small Group Dynamics and Productivity in the BSCS Laboratory Block Program," Journal of Research in Science Teaching, 4: 67-73, 1966.

Walberg, H. J. and A. Ahlgren. "A Social Psychological Model of the Classroom." Paper presented at AERA Convention, Minneapolis, Minn., 1970.

Walberg, H. J. and G. Anderson. "Classroom Climate and Individual Learning," Journal of Educational Psychology, 59(6): 414-419, 1968.

Shymansky, James A. "How Is Student Performance Affected by the One-to-One Teacher-Student Interactions Occurring in an Activity-Centered Science Classroom?" Journal of Research in Science Teaching, 13(3): 253-258, 1976.

Descriptors--*Activity Learning; *Educational Research; Elementary Education; *Elementary School Science; *Instruction; Science Education; Student Behavior; *Student Teacher Relationship; Teacher Role

Expanded Abstract and Analysis Prepared Especially for I.S.E. by William R. Brown, Old Dominion University.

Purpose

This correlational study dealt with the relationships between one-to-one interactions taking place between a teacher and a student in an activity-centered science classroom and that student's behavior during the remainder of the science lesson.

Rationale

Shymansky summarized the literature concerning classroom interaction analysis techniques and generalized to produce this rationale. Individual differences among students can affect both the actual and perceived interactions in a classroom. Intraclass variations in teacher behaviors do occur. Therefore, it is appropriate to investigate dyadic teacher-student interaction. Dyadic interactions are especially important when a teacher is expected to interact with children long enough to guide the child's behavior but not so long as to stifle hands-on investigation by the child.

Research Design and Procedures

The variates were length of interaction (time interval) and the nature of the interaction (verbal or nonverbal). The criterion variable was the student's classroom behavior. This variable was sudivided into ten categories.

L0	Miscellaneous
L1	Observes demonstration
L2	Follows teacher's directions

L3 Does activity of own design
L4 Responds to teacher
L5 Initiates interaction with teacher
L6 Initiates interaction with another student
L7 Receives ideas from another student
L8 Copies other students
L9 Gives ideas to another student

The revised SCAS Classroom Interaction Categories--Student Behaviors was used by nine trained observers in fifth-grade science. All data were collected using this instrument. The intra- and interobserver reliability coefficients of the observer team were calculated according to the Scott formula. All coefficients exceeded 0.75.

A total of 78 student observations were made over a five-week period. Fifth graders were randomly selected from a class of 26 students. Each student was then observed by a single observer for an entire class period. The observer coded the student's behavior every three seconds or when the behavior changed in less than three seconds. Nonverbal one-to-one teacher-student interactions were recorded as "O." Verbal interactions were coded "V." When the teacher interacted directly with the student being observed, a three-symbol code such as OL3 was recorded. This represented nonverbal, student does activity of own design. VL1 would be verbal, student observes demonstration. A two-symbol code was used when the teacher-student were not interacting directly.

Data for the variates were summarized by three scores. The V-score is the percentage of observational time the teacher verbally interacted with a student in a one-to-one situation. V is equal to the sum of codes containing V divided by the sum of all codes. The O-score is equal to the sum of codes containing O divided by the sum of all codes. The O-score is the percentage of observational time the teacher interacted nonverbally with a student in a one-to-one situation. The percentage of observational time the teacher interacted verbally and nonverbally with a student in a one-to-one situatior (all dyadic observations) is the O + V score.

The criterion variable (Student Classroom Behaviors) was represented by ten categories plus an L-score. L is equal to the sum of all codes

containing L divided by the sum of all codes. Student behaviors coded during the one-to-one verbal interaction with the teacher were not included in the L-scores. It was decided that what the student did in the classroom when not under the direct influence of the teacher was far more important than those behaviors exhibited during the interaction.

The relationships between variables were analyzed using Spearman rank-correlation.

Findings

The total amount of time that the teacher spent interacting directly with the student (O and V) correlated negatively with the total amount of time that the student spent doing activities of his own design (L3).

Positive correlations (significant at the $\alpha = .05$) were:

L1	O & V	V	
L4	O & V	V	
L5	O & V	V	O

Interpretations

Two explanations were presented to interpret the negative correlation. Perhaps increased teacher interaction with the student reduced the amount of time available to the student to exhibit any L3 behavior. A second explanation suggests that the teacher tended to interact longer with students who were not involved in productive activity. Both of these explanations are discounted by the investigator based on the patterns in the data.

The positive correlations for L1 and L5 indicate that students with whom the teacher interacted for longer periods of time tended to be very dependent on the teacher. The lengthy one-to-one teacher interactions may have produced insecure feelings in students regarding their performance in the activities. One-to-one interaction may result in interference if the interaction proceeds too long. Is there an optimum

interaction time between teacher and student when the student is involved in a hands-on activity-oriented science experience? The investigation suggests that teachers must share the teaching with materials, thereby reducing the teacher-to-student interactions when a child is manipulating the materials or engaging in independent activity.

ABSTRACTOR'S ANALYSIS

This relatively simple, straightforward study identifies a potential experimental situation. It would be possible to manipulate the variates in Shymansky's study to attempt to establish cause and effect. The length and nature of the teacher-student interactions could be systematically manipulated. Does the length and/or nature of the teacher-student interaction affect the "security feelings" of students involved in hands-on activities? Does a "long" student-teacher interaction interfere with student's involvement in hands-on activity-centered science?

A few questions can be raised concerning information presented in the report. The selection and description of samples appears to be a perpetual problem in report writing. Fifth-grade science classes were used in the study. How many classes were used? How many teachers were used? Seventy-eight interaction sessions with one teacher and 78 different students would most likely yield different results from 78 observations using a dozen different teachers. Why were fifth graders used? No description of the population was presented. Did the teachers observed have a history of student-teacher interactions? Did the students observed have a history of teacher-student interactions? What science program was being used? Was it always the same lesson? What control was kept to insure that each of the 78 episodes involved as equal amount of expected hands-on activity? What procedure was used to randomly select a student from a class? Did the selected student know he was being observed? What precautions were taken to prevent observer-subject interactions?

Reliability indices were presented for the instrument used. No measures of validity were given. Are the nine behaviors listed on the SCAS the student behaviors we should be examining?

The overall length of the study was five weeks, but each teacher-student observation was 30-40 minutes. What might be the patterns of teacher-student interactions if there were several observations made of the same teacher-student pairs by the same or by different observers? Perhaps additional clues to what is happening in a classroom could be retrieved if the sessions were recorded on video tape using a two-camera system, one for the teacher and one for the student.

Two explanations were presented concerning the negative correlation between O & V and L3. Reasonable theses were given to discount these two explanations. Unfortunately, an alternative interpretation was not given. What might this negative correlation mean?

The investigator suggests that teachers must share the teaching with materials. In the last couple of years, activity cards have been written for use with predominantly hands-on science programs (1, 2). These cards can serve as a method to reduce teacher-student interaction and still have a student involved in a hands-on, inquiry situation (3). Short reading materials can also be used to decrease direct teacher-student confrontation (4).

This investigation points to variables that may be examined to provide insight into the cause-effect relationships occurring in activity-centered classrooms. We need to continue this line of research and combine similar studies in order to develop instructional principles.

REFERENCES

1. White, Edwin and William Brown. Delta Project Cards for ESS-Elementary Science Study, 14 titles. Nashua, NH: Delta Education, Inc., 1978.

2. Brown, William and Edwin White. OBSERVE Activity Cards for ESS, 6 titles. Virginia Beach, VA: OBSERVE, 1979.

3. Martin, Jean. The Effects of OBSERVE Activity Cards on Achievement and Attitudes of Sixth Graders Using ESS-Batteries and Bulbs. Unpublished master's problem paper, Old Dominion University, Norfolk, VA, 1979.

4. Brown, Betty, William Brown, and Edwin White. Readers for ESS, 10 titles. Virginia Beach, VA: OBSERVE, 1979.

Spears, Jacqueline and Dean Zollman. "The Influence of Structured Versus Unstructured Laboratory on Students' Understanding the Process of Science." Journal of Research in Science Teaching, 14(1): 33-38, 1977.

Descriptors--College Science; *Educational Research; *Higher Education; *Inquiry Training; *Instruction; Physics; Science Education; Science Laboratories

Expanded Abstract and Analysis Prepared Especially for I.S.E. by Marvin F. Wideen, Simon Fraser University, Vancouver, Canada.

Purpose

The purpose of the study was to determine the relative effectiveness of structured and unstructured laboratories in teaching an understanding of the process of science.

Rationale

This study falls into the category of those conducted over the last few years contrasting various types of structured and unstructured teaching strategies in post secondary science teaching. While previous studies of this type have compared the two strategies in terms of their effectiveness in teaching concepts, developing formal operational thinking, or aiding the development of attitudes, the study described in this paper used understanding of the process of science as a dependent variable. The authors argue that the structured laboratory is more in keeping with the position of Gagné and Piaget, both of whom suggest that some form of structure is needed. The unstructured situation, on the other hand, is more in line with the theoretical arguments of Bruner who argues for practice in inquiry. In effect, the study is a contrast of these two positions.

Research Design and Procedures

A pretest, posttest control group design was used with laboratory structure as the independent variable and student scores on the Welch

Science Process Inventory (SPI) as the dependent variable. The latter was seen as a measure of the students' understanding of the process of science. The treatments included two approaches to structuring the laboratory section of a course, Man's Physics World I. One included the traditional format in which the student was given instructions to help verify principles in class; the second included an approach in which the student discovers or inquires into principles discussed in lectures. The lectures, which were presented by four different individuals, were not part of the experiment. Pretest scores on SPI, laboratory grades and lecture instructor were used as covariates to adjust pretest scores.

Instrument

The dependent variable, students' understanding of the process of science, was measured using the Welch Science Process Inventory (SPI). This 135 statement rating scale uses a dichotomous response format. A reliability estimate of .86 (Kuder-Richardson) is reported as well as a statement that "validity has been established through both predictive and constructive measures." The instrument provides both a total score and scores on four subscales: assumptions, activities, nature of outcomes, and ethics and goals. The instrument was administered during the first week of the semester and as a posttest during the last week of the course.

Sample

The subjects were freshmen and sophomore students from areas such as elementary education, business administration, home economics and the social sciences who were taking the introductory physics course to satisfy a general science requirement. The majority of the students had had exposure to biology or geology as a college science but only 23 percent had completed high school physics. The subjects were assigned at random both to the four lecture groups and to the laboratory sections.

Findings

The adjusted posttest scores indicated that no differences occurred in three of the components: assumptions, nature of outcomes or ethics, and goals. In the case of activities, however, students in the structured laboratory scored higher.

The authors point out that those in the structured laboratory were led through activities performed by scientists while those in the unstructured laboratory, in the absence of specific instructions, turned to lecture notes and textbooks for explanations of observations. The authors further argue that this is typical of the concrete-operational student and implies that were more students of a formal operational level, they would have performed differently by making observations, building models and testing them. The authors conclude that structure, because it provided examples of activities, caused the students to better learn the process of science. The authors further conclude that the "...average college freshman or sophomore taking his first physics class apparently requires a structured experience and training in scientific process before he will understand it" (page 38).

Interpretation

The study compares two strategies to teaching laboratory physics, structured and unstructured. The authors describe their study as an extension of work done by Gunsch (1972), Murphy (1976) and Tanner (1967), who similarly compared these two strategies but using different dependent variables. The authors argue that their study, which uses the understanding of the process of science as a dependent variable, provide a useful contribution. Indeed, this would be the case were all studies dealing with the same treatment. Closer examination of the studies cited by the authors reveal that this is not the case. Gunsch compared laboratory and nonlaboratory approaches to teaching science and Murphy compared process and content centered laboratories. The Tanner paper was not a comparative study at all but a description of two approaches. Thus the study seems unrelated to a matrix of previous studies and as such it is difficult to assess its conceptual contribution. At best, the study should be seen

as an evaluation of two laboratory approaches with no justifiable basis for generalizing to previous research in the area.

Similar difficulties exist when the authors try to draw links to the work of Bruner, Piaget and Gagné. It is not clear from the paper and, indeed, highly questionable that the two strategies used are a reflection of the positions of any of these three psychologists. These observations should not necessarily be seen as a criticism of the authors as much as a criticism of the state of current research with regard to the use of laboratory work in science teaching at the college level. The authors can hardly be held responsible for the fact that there are probably an insufficient number of studies dealing with laboratory teaching which are sufficiently similar to create a matrix from which to draw a conceptual framework.

The authors followed a fairly standard methodological paradigm in conducting the study. Two general treatments were compared using a non-equivalent control group design. Later, I will comment on some ways that this specific study could have been improved. First, however, I should like to comment on the basic weakness in comparative studies of this type. Comparing general treatments such as structured and unstructured approaches to teaching laboratory science always raises the question: What is it that is being compared? When differences do occur it is difficult to know how to account for such differences given the range of things that might occur under either treatment.

The authors are to be commended in their use of the analysis of covariance and, in particular, the selection of variables to control for past experience. However, the context within which the study was conducted created a set of problems typical of studies of this type. The authors were working within the context of a fairly structured, traditional course involving lectures and laboratory. It would seem also that the outcomes of the course and the students' expectations would be similarly traditional. It is into this context, then, that the authors introduce an alternative approach to laboratory teaching, an inquiry approach. But it is still likely required of the students that they perform in traditional ways on the final examination. Thus, it is not

hard to understand how the unstructured treatment became something quite different from the steps the authors describe on page 35, Table 1, which implies having students solve problems and draw conclusions to students citing "...lecturers and lecture textbooks for their explanations..." (page 37). In short, can one fairly test the merits of two alternative strategies in a setting that so strongly favors one over the other?

The design of the study could have been strengthened had the authors a) used a wider range of criterion measures, b) used a factorial design, and c) established a more logical link between the treatment and the criterion measure. The discussion turns to a brief comment of each.

The use of a single criterion measure in a study of this type is tantamount to placing all one's eggs in one basket. The study could have been improved significantly if the authors would have used measures of content, attitude and interest. From these one could draw more inferences about the relative effectiveness of the two approaches. Second, the use of a factorial design in which selected attributes or demographic characteristics of the subjects were crossed with the treatments would have made it possible to identify different groups for whom the treatments may have been more effective. Third, the argument that one of the treatments is likely to be more effective than the other in improving the understanding of science is lacking in the paper. The rational links between the dependent and independent variables are simply not there.

While the sample was well described, the paper should have included the number of students who were involved in the study and provided a more explicit explanation as to why the attrition rate was as high as 50 percent. Also, it is disconcerting when researchers do not include standard deviations in tables that have means and F-values. The conclusions seemed rather unwarranted. The authors concluded that "...the average college freshman or sophomore taking his first physics class apparently requires a structured experience..." on the basis of a 1.5 difference on adjusted means on one subscale when three other subscales showed no significant difference. An alternative hypothesis would be that the unstructured approach to laboratory work was equally as effective as the structured approach since three of the four subscales showed no significant differences.

Earlier, it was mentioned that research in the use of the laboratory in science teaching is limited. Future research might well take two directions. First, it would be useful to have reports of systematic observational studies which describe what goes on in laboratories under the aegis of laboratory teaching. These should be designed to describe the transactions which occur in laboratories in order that specific variables can be identified for future research. But such studies should also deal with the contextual, social and political attributes of teaching. With such knowledge, one can then more appropriately design comparative studies based on specific variables rather than global treatments. Developmental research also needs to be done to provide a conceptual basis for what is done in laboratories, based on the work of eminent psychologists such as Piaget, Bruner and Gagné. With a basis formed by such activity, then comparative studies such as the one just reviewed can take a more refined focus and contribute to a more effective conceptual base.

REFERENCES

Gunsch, L. M. "A Comparison of Students' Achievement and Attitude Changes Resulting from a Laboratory and Nonlaboratory Approach to General Education Physical Science Courses." Unpublished Ed.D. dissertation, University of Northern Colorado, 1972.

Murphy, G. W. "A Study of the Relative Effectiveness of Content and Process Centered Biology Laboratories for College Freshmen." Unpublished Ed.D. dissertation, University of Kentucky, 1967.

Tanner, R. T. "Expository-Deductive Versus Discovery Inductive Programming on Physical Science Principles." Journal of Research in Science Teaching, 6: 136-142, 1969

Sunal, Dennis W. "Analysis of Research on the Educational Uses of a Planetarium." Journal of Research in Science Teaching, 13(4): 345-349, 1976.

Descriptors--*Astronomy; *Educational Objectives; Educational Research; *Evaluation; *Instruction; *Planetariums; *Science Education; Science Facilities

Expanded Abstract and Analysis Prepared Especially for I.S.E. by Gerald H. Krockover, Purdue University.

Purpose

The purpose of Sunal's study was to "analyze the development and use of a model for evaluating student outcomes involving a school-associated planetarium" (p. 345).

Rationale

The rationale for this study was to identify "what is known through research as to the actual role of the school and college-associated planetarium in education" (p. 345).

Research Design and Procedure

A model was developed for defining the school-associated planetarium, including its perceived role in education. The procedure involved using the model to analyze past research studies including planetarium studies and test data, categorization of test instruments, analysis of test data, and evaluation of results.

Nine out of a total of 16 studies concerning the planetarium were isolated as possibly providing information related to analysis of the model presented. The selected studies involved more than 4000 students from grade two through university freshman. The type of planetarium and number of visits also varied between studies.

The analysis of the data involved identical computer statistical analysis and significance levels of student data as performed in the

original research report. The difference involved analysis on question subgroups rather than on total score for each student. The test data were analyzed on three levels: elementary, secondary, and college. This resulted in four independent sets of data relating to the elementary level, three at the secondary level, and two at the college level. The studies involved single and multiple classroom and planetarium experiences using a variety of comparisons. The goal areas analyzed included three cognitive levels and four process skills.

Findings

Research so far does not support the belief that a single planetarium lesson produced better results, in any goal area, than a single classroom lesson covering the same objectives.

Classroom instructional units involving one to eight planetarium lessons as compared to units with no planetarium involvement showed no significant difference at the elementary levels and mixed results at the secondary and college levels.

Practices reported which produced better results than typical one-visit usage were orientation sessions given before planetarium visits, use of the planetarium for remedial lessons, coordinated and combined planetarium classroom instructional units, and visits scheduled late in the classroom unit. However, gains affected by the planetarium experience were found not to be influenced by IQ, reading or mathematics ability of the student, sex of the student; or preplanetarium and follow-up exercises.

Interpretations

Research in planetarium education has tended to be specific, dealing with few aspects of what planetarium educators define as goals. In the years ahead, planetarium education must become a far more systematized process with objectives better stated in terms of concrete, observable, and trainable behaviors and should include the establishment

of a relationship between research and actual decision making. Research to date has been more successful in challenging old assumptions than it has been in establishing new generalizations.

ABSTRACTOR'S ANALYSIS

This study can be summarized in one word: "confusing." The author purports to use a "model," yet he never defines what the model is. The author submits a classification chart for goal areas of the model without ever stating the model. One questions whether this article is really an analysis of research or simply a summary of previous surveys. Many questions can be asked about this article including:

1. What was the basis used for the selection of the model?
2. What specific procedure was used for using the model to analyze past research studies? (p. 346)
3. Why did the author cover grades two through college? (p. 346)
4. What statistical analysis was used and why wasn't the level of significance reported? (p. 347)
5. Why did the author use subgroup data when the original researchers used total score data for each student? (p. 347)
6. Why did the author use fifteen variables? (p. 347)

In conclusion, Sunal states five implications derived from the "operational school planetarium model," which may be the best part of this article. Unfortunately the implications cannot be traced directly to any parts of the study.

RESPONSE TO ANALYSIS

IN RESPONSE TO THE ANALYSIS OF

Mills, Terence and Eubanks, I. Dwaine. "Student Assessment of ISCS Classrooms and Grade Placement of ISCS Level I and II," by Dorothy Gabel. Investigations in Science Education, 5(4): 19-22, 1979.

by
Terence Mills
Oklahoma State University

The major purpose of the study under review was to determine the opinions of students concerning their ISCS classrooms and teachers where the schools interchanged ISCS Levels I and II within the junior high school grades. Information was gathered to determine if differences in student opinion existed between (1) seventh and eighth-grade students using Level I materials, (2) eighth and ninth-grade students using Level II materials, and (3) eighth-grade students using Level I and eighth-grade students using Level II. An additional objective was to obtain a general assessment of student opinions concerning their ISCS classroom.

Questions asked and information requested in the critique follow as does rationale and data which clarify the strengths and weaknesses of the design and outcome of the study.

1. *Were students randomly selected from all school systems using ISCS in the manner described in the article or were students selected from one state or from one local school district?*

   The 1,967 subjects were students from four school districts. The population was not a random sample drawn from the 1,967 students. The wording ". . . a sample of 1,967 Junior High School students in 75 sections of ISCS science" is misleading. The term "population" rather than "sample" would be a better choice.

2. *Are students in urban, suburban, and rural environments represented?*

One urban and three rural communities in northeastern Oklahoma were represented. The majority of the subjects were from the urban district, with 178 or 9 percent of the total population coming from three rural communities. Urban and rural student responses were compared but not reported as part of the study.

3. *Are there any reliability coefficients for the assessment instrument?*

Reliability coefficients for the 20 assessment items (Your Student's Role, 1973) were not available and none were calculated. The fact that the students' perceptions were obtained at different times, dates, grade levels, communities, and schools by different teachers of a large number of different students in a self-paced curriculum made a quantitative measure of reliability desirable but difficult to establish. In this case validity was considered more important than reliability as it is possible for a test to be sufficiently valid for practical purposes without being very reliable (Guilford, 1956). Content and face validity were determined by trained ISCS personnel. Validity was also inferred from the fact that those who developed the objectives and materials for the ISCS curriculum accepted and published the assessment instrument for use in ISCS classrooms. It was understood that a reliability coefficient of .95 for each of the 20 items would have enhanced the survey.

4. *Information on the sample size of each group and how students were clustered under teachers was not reported.*

Sample size of each group was available upon request as was noted at the bottom of page 321, i.e., "*Summaries of actual and expected student responses to the questions are available upon request." Sample size was as follows:

	N
Seventh Grade - Level I	472
Eighth Grade - Level I	568
Eighth Grade - Level II	491
Ninth Grade - Level II	436

To protect the anonymity of students and teachers it was requested that the response sheets not be coded. It was felt that the voluntary participation of teachers required that responses not be traced to teachers. Therefore, analysis of student response by teacher was not conducted.

5. *"Because a large number of students were surveyed in this study, small differences in means produce statistical significances. Examination of means in Table II shows that differences of 0.2 on the 5-point scale produces statistical significance."* (Gabel)

No statistical comparison of means were made in this study! The Chi-square one-sample test was used to determine if a statistically significant difference existed in the proportion of responses to each of the items when grade or ISCS level varied. Mean response to items were presented only as means. Comparisons, other than a "generally favorable response to the ISCS curriculum materials and classroom teachers," were not made by the authors. It was suggested that the positive mean response by group implied support for ISCS materials used out of grade.

6. *Was a control group used for comparison?*

Fourteen of the 19 teachers cooperating in the survey also taught other junior high school science courses. Data from 471 students in 16 non-ISCS science sections was gathered using the same assessment instrument. The implication that "ISCS students generally rated the (ISCS) course positively" (Mills and Eubanks, 1979) is strengthened by comparison of the ISCS and non-ISCS mean responses in Table I. For all positively stated items except 6 and 10, the ISCS student means were higher than non-ISCS student means. The lows means for the negatively stated item (1, 13 and 15) indicate a more positive response from ISCS students on Items 1 and 15. A significant difference in ISCS and non-ISCS response frequencies occurred in 13 of the 20 items.

Table I. Means and Chi-Square Values for ISCS Questionnaire

Item	ISCS (N=1,967) $\bar{X}$	ISCS (N=1,967) S.D.	Non-ISCS (N=471) $\bar{X}$	Non-ISCS (N=471) S.D.	X^2 Values X^2
1. My teacher expects students to agree with his views.	2.98	1.5716	3.01	1.5658	26.22*
2. My teacher seems to be interested in teaching.	4.08	1.8249	3.74	1.7591	125.03*
3. My teacher cares whether I learn.	3.95	1.8065	3.59	1.7305	63.36*
4. I am interested in the subject matter in this class.	3.32	1.6721	3.23	1.6489	3.71
5. I enjoy the activities in this course.	3.42	1.6905	3.28	1.6528	14.97*
6. I believe this course challenges me mentally.	3.56	1.7177	3.56	1.7188	0.38
7. Our teacher encourages us to express our opinions.	3.51	1.7007	3.44	1.6860	7.44
8. Our teacher seems to accept new ideas and viewpoints.	3.54	1.7048	3.37	1.6669	15.54*
9. We are encouraged to ask questions.	3.76	1.7615	3.68	1.741(	5.10
10. The course is worthwhile.	3.47	1.7075	3.48	1.7052	3.09
11. The course is enjoyable most of the time.	3.46	1.6977	3.20	1.6425	27.46*
12. When the teacher talks to the class, I listen.	3.48	1.6801	3.23	1.6243	24.47*
13. My teacher gets most of his opinions about me from my answers to his test questions.	2.86	1.5446	2.85	1.5313	11.89*
14. Our teacher lets us make some decisions about how the class will be run.	2.47	1.4665	2.37	1.4123	11.37*
15. My teacher makes me feel stupid when I make a mistake.	2.47	1.4878	2.51	1.4734	17.90*
16. Equipment and supplies are easy to get when I need them.	3.87	1.7880	3.61	1.7233	32.09*
17. Our teacher likes to learn.	3.72	1.7401	3.50	1.6893	36.68*
18. We are encouraged to solve problems.	3.97	1.7963	3.92	1.7909	3.32
19. I know where the safety equipment is located.	3.51	1.7327	3.27	1.6792	16.14*
20. I feel that the equipment and supplies belong to the teacher and the students equally.	3.65	1.7622	3.55	1.7379	9.20

*X^2 values > 9.49 indicate differences in response frequencies which are significant at the 0.05 level or greater (D_f = 4 in all cases).

Item number 14 was an exception in that at the request of the participating teachers, student responses by teacher were tabulated. The low mean response was consistent for all sections of both ISCS and non-ISCS classrooms.

A number of compromises were made in an attempt to gather data from as large a population as possible. Recognizing the limitations of such a study, the authors made no attempt to generalize to other populations. The information was of value to participating school personnel making curriculum decisions. It is hoped that the additional information in this reply to the analysis will make the study of greater value to curriculum decision makers.